终端用户行为研究的实地调查经验

顾立平

飞资得系统科技股份有限公司　台北　100140

[摘要] 以对研究伦理和研究态度的强调贯穿全文，简介 2006-2009 年间在完成博士论文《网络信息行为的差异化用户模型——创建与使用型人》前后所做的 6 次调查经验，包括研究思路与策略、研究角度与立场、研究设计的纲要、研究对象的特殊价值、抽样与调查的审核过程、合理推论原则以及整理 SPSS 工具的操作步骤等。

[关键词] 用户行为　抽样调查　假设检验　合理竞替　方法论　研究伦理

1　研究思路与策略

目前图书馆学主要面临的是网络带来的冲击，大量新型的网络技术对数字图书馆和图书馆信息服务带来了巨大挑战，对数字图书馆技术方面的研究造成决定性影响；伴随网络技术的革新，海量信息的处理、分析与计算成为情报计量、书目计量和科学计量方面的新课题；信息组织、信息检索与信息需求的改变，促使了用户行为的改变，脱离了传统图书馆学的假设；最后，这些改变与冲击的加总，就造成了图书馆性质的根本性变化，从而使图书馆学有了新的方向与新的课题。所以，现在研究图书情报学，可谓身逢其时。

然而，现代的信息情报分析日益复杂、图书馆服务日益增加，数字图书馆的建设涉及到很多具备不同技能的参与者。这就对知识服务者的学科背景以及研究工作提出了很多新的要求和挑战。专业人员必须拥有广泛学科的知识基础，方能从宽广视野中对实际从事的图书情报工作产生独特的见解与创见，还必须从广泛的、有新意的甚至天马行空的想法上，借由文献阅读与讨论，通过与实际工作相结合，并经过严谨的科学研究过程，逐步聚集到两三个具体的前沿问题上，在钻研、攻关、克服种种困难后，提出创新与贡献。

以终端用户（end-user）行为研究而言，有两种基本思路：①从图书馆学出发，结合其他学科；②以其他学科如管理学、社会学、心理学、经济学等的实践基础，作为指导者或者引入而成为解决图书情报事业发展困难的理论来源，利用其他学科的内容来建立假设，用统计或者其他方法来进行实证调查，解决当前图书馆面临的某个问题，获得可靠的结论。

刚开展的课题研究，总是困难重重，执行之前，有能够想象得到的困难，执行过程中，会出现当初没有想象到的困难。甚至课题完成后，才会发现真正重要的问题。本文介绍的《网络信息行为的差异化用户模型——创建与使用型人》（以下简称《型人》）的研究经验，先是从信息搜寻行为出发，建立假设，以反证求取基本行为描述，在进行研究解释时，又从心理学与社会学的若干理论建立合理竞替假设，进行研究推论，从而得到一个假说。在此我们并不着墨《型人》，而是概述当时实地调查的方法论，以期抛砖引玉。

2　研究角度与立场

当代研究论文的写作规范和要素，包括找到什么问题、提出什么问题、解决什么问题以及用什么方法、用什么技术、用什么工具来解决这些问题。然而，自 2009 年至今担任 *International Journal of Library and Information Science* 和 *International Journal of Science and Technology Education Research*

编委以来，经常审阅到大量南亚和东南亚论文，其中，最后被退稿的论文大多是立意不够鲜明，而不是缺乏以上要素。在众多论文评审意见中，最常见的是“请明确您的研究价值与意义”。换言之，我们希望研究者们挑战自己的心志、要求自己、苛求自己，去贴近实地现况，理清自己的研究目的和方向。简言之，质朴远胜于取巧。

此外，如果是终端用户行为研究，我们还注意研究报告所透露的细节，审核研究者是否踏入伦理与政治的误区。最基本的要求有如下几方面：

- 自愿参加原则：不强迫受访者参与，所有参与研究的受访者都是自愿参与。
- 对受访者无害原则：研究内容不涉及受访者个人评价、隐私以及研究单位的态度。
- 匿名和保密原则：对所有受访者数据，采用所罗门加密法，并且予以安全存放保密。
- 研究过程透明原则：研究者公开姓名与单位，研究问题不涉及单位对个别受访者的特殊照顾与福利政策。
- 研究报告中立原则：研究者在研究报告中具体指出研究本身存在的缺点，而不加以掩饰。

不过，有些内容是审核者不会看到，却又十分重要的方面。例如，研究本身以外的另一种实际问题是：研究者如何面对自己以及所属机构的既得利益？因为科学研究本身也属于社会文化的一种实践，而这不可避免地涉及到以下问题：用什么角度？关心什么问题？意图得到什么结论？每个人在一项特别的研究事业的结果中都有既得利益。研究者如果清楚地意识到自己的既得利益，以及研究项目中各方复杂的既得利益后，应该让这些既得利益不影响研究设计，并且防范研究结果被刻意地、有意地或无意地误导。

事实上，只要能够想清楚上述要点，人们很容易从自己的研究问题以及别人的研究成果中找到切入点和突破点。研究伦理不是一种道德说教，而是所有研究最重要的关键所在。如果说，研究者的研究伦理决定了其研究高度和贡献，一点也不为过。其他的研究难点其实都可以解决，例如不会 SPSS 可以请人指导，研究设计不当可以重做或者补救，研究推论不佳可以改善，唯有研究伦理存在于个人心中，贯穿整个研究过程，不可不慎。

研究者应该保持一个知识分子该有的良知，对“政策导致结论”应有所警惕。几年前，当《型人》还未成形，而只是知道要进行某些社会科学调查前，指导者和研究者就共同面临过“我在为谁研究”的问题。这个简单的自我提问，对于日后形成《型人》却具有决定性影响。

3　研究设计的纲要

当代的研究人员，不再只从事理论综述和著述，由于图书情报工作的发达，几乎所有学科（至少是中科院的理工类学科）都把“综述”当成“必须要做”而非“唯一要做”的事情，并且把“创新”看作“必须能做”的决胜条件。

几年前，经过仔细研究，我们认为，无论从哪个子领域的某个专题开始着手，创新过程都仍然不离以下三方面：①文献综述——发现尚未解决的问题——用旧方法解决新问题。②研究方法——发现旧方法的不足——产生新方法。③研究结果——归纳新方法、新问题的研究结果——产生新领域。三者环环相扣，形成创新型研究。特别注意的是：对新问题、新方法、新领域的提出，必须建立在实证的基础上，可以被其他研究者重复，或者进行修改。

以《型人》的研究设计为例，最初设计执行 6 次相同问卷调查的纵向研究，后来发现许多执行面的问题，于是改采 6 次不同问卷调查的横向研究，研究计划实施到一半，发现可以执行，但是无法深入描述终端用户实际的工作内容以及网络行为，于是加入了访谈法和自然观察法，改用质量混合型研究。

在进行总体设计时，可以按情况调整，但是实际执行阶段，则容不得随意更改计划；所以，为确保研究计划“如期依约”完成，在进行研究设计的时候，参考了许多前辈的研究报告和教科书，制定一些硬性

规则。例如，就问卷调查和访谈调查的硬性过程来说，有以下要点：

- 问卷调查的具体步骤包括：①确定问卷框架和主要内容；②设计问卷的具体问题；③在小范围内试填问卷；④根据试填问卷中暴露出的问题修改问卷；⑤设计问卷发放与回收方式；⑥发放问卷；⑦被调查者回答问卷；⑧回收问卷；⑨追踪未回答问卷；⑩整理问卷，剔除作废问卷；⑪对问卷答案进行编码，以便处理分析。
- 访谈调查的具体步骤包括：①确定访谈主要内容；②确定访谈基本形式；③设计访谈提纲；④安排访谈过程；⑤实施访谈；⑥整理访谈记录；⑦回忆访谈情况并再次整理记录；⑧对访谈结果进行汇总，以便处理分析。

由于充分考虑研究伦理的每个细节，并且拥有坚定而且明确的研究态度，所以在“师出有名”的情况下，能够准确评估研究设计的可行性以及实际执行的效率。

4 研究对象的特殊价值

在执行用户研究时，需要非常清楚谁是研究对象、谁不是研究对象，否则无法进行之后的研究推论。以《型人》为例，当初选择中国科学院研究生院的硕、博士研究生作为研究对象，是因为其具有特殊性，表现在：与他们的上一代科研人员不同，这一代中国科学院所有基本的人口统计参数、社会参数、经济参数以及网络通讯参数，对比过去 15 年都有新的数值，从事研究的科学研究人员总人数呈指数增长，在中科院内的比率大幅增加。此外，在研究生院的学籍统筹下，分属到直属中国科学院的各个研究机构，他们在研究所里的情况是这样：与正式工作的职员一起，通过实验室或个别指导，将技术、方法、理论、研究规范和价值观传递出去。这些身处工作岗位的前辈，不仅把他们的技术、方法、信息和理论传授给和他们一起工作的新手，而且也把指导研究的研究所规范和价值观传给了其同事。

研究上述对象，不容易得出普遍性的用户行为，但是其样本代表性特殊而且特征集中，容易找到有趣（interesting）或有新意的发现；但是另一方面，也很具有挑战性，容易产生设计不良或者操作不良的问题。克服这个难点的方式，就是反复仔细考虑研究伦理与态度。在研究过程中，几乎每天都在天人交战：来自用户的直接冲击，会不断修正研究者已有的知识结构以及“普通常识”。这种锻炼对于日后从事其他研究课题，甚至在不同图书馆以及公司工作，都是很有帮助的。另一方面，尽管面对的读者群不同，但是《型人》的研究目的是提升图书馆的系统与服务而非研究者发文数与被引量，研究立场是读者立场而非图书馆长与馆员的立场，研究利益是读者利益而非图书馆长与研究者的利益。所以，当研究者日后面对来自各种自称“代表用户”的声音时，能够直接参照几年前对终端用户的实地调查经验，直觉地发现对方的潜藏用意以及事务上隐而未现的危机。在旁人看来，好似一种特异功能，其实只是几年前下过苦功，而且每隔一段时期，花一小部分时间去温习而已。

5 抽样与调查的审核过程

抽样似乎是很复杂与困难的工作。但是，如果掌握研究伦理与态度的原则，并且长存在心，那么考虑如何抽样并不是太困难的事情。虽然有很多种抽样方法、抽样技术和抽样量表，但是如果不符合研究伦理与态度，那有何用？反过来说，即使不懂得抽样，只要把握研究伦理与态度，可以自己推导较为简便的抽样方式。如果只有重视抽样技巧而忽视其研究伦理，则最后必然产生戏剧性的研究溃败，不可不慎。下面列举《型人》研究前的思考过程。

5.1 母体样本名册是否可得?

随机抽样与方便取样，在操作面上的最大不同，就在于随机抽样是在所有调查对象（母体）中选择受访者（样本）。因此，没有包括所有母体信息的样本名册，就不可能进行随机抽样。在这个阶段，除了提出明确的研究计划书以外，也对于信息保密作出承诺与具体措施。

5.2 母体样本名册是否完整?

拥有名册并不等于掌握了母体的相关信息，对于问卷调查而言，名册至少必须记载以下信息：姓名、就读系所、是否就读于北京地区、是否仍在就读。上述信息提供研究者判断是否能够进行随机抽样。事实上，该名册包括个人信息，如性别、攻读学位、年级或入学时间、年龄等，但是不包括他们的联络方式。研究者在这个阶段开始判断采用哪种类型的随机抽样。

5.3 如何进行抽样?

仔细考虑后的《型人》采用分层随机抽样，对各个京区研究所的学生比例、性别比例、博硕士比例、年级比例等进行抽样。从每一个抽样单位中，按照这个抽样单位的规模成比例地进行选择。其优点是：①能够保证形成单位分类基础之特征的代表性。因此，所产生的变异性比简单随机抽样或多阶段随机抽样更少；②通过分类过程可以减少总体成员未能被包括进来的机会；③能够估计每一层的特征，因而可以进行比较。但是，其缺点是：存在着错误分类的可能性，因此会增加变异性。为了克服这个缺陷，进行过变异系数的短暂研究，最后虽然无法求证可能性，但是意外地学会结构方程式。也因为如此，在访谈调查上，就采用目的抽样。以问卷调查的分析结果为基础，通过判断，从总体中选择一个子群体作为整体的代表，将这一群体中的一部分作为样本。

5.4 如何有效发放问卷?

①与时间、地点和人力有关。发放时间（时期）越长，地点越分散，越是有效，但是人力协助的要求也越高。②与人力、礼品和经费有关。人力越多，礼品越多，发放问卷的数量越多；研究效度就越能增加；研究效度越是增加，越有能力牺牲效度换取信度，从而解决抽样所带来的信度不足。③与问卷数量、问题数量、礼品价值和经费有关。问卷数量越多，问题数量就应该越少；这个问题在统计学上有一套公式，可以换算出参照值，在《型人》完成之后才知道这个解法，但是《型人》也确实在“安全区域”内符合标准；不过，若从实践角度而言，与其考虑问卷数量和问卷题目，不如更多地考虑行政因素和受访者预期心理所带来的回答偏差，从而造成信度大幅降低的问题。

5.5 样本是否愿意回答?

是否愿意回答包括两个问题：①调查如何进行？也就是上述的发放问卷问题。②问卷本身是否不造成受访者焦虑。造成受访者焦虑的原因很多，除了发放回收的过程，最重要的就是问卷是否能够轻易作答。比如，不适合问太新的议题，比如近期才出现的网络词汇。

如果一定要问最新的网络行为，那么就要对是否了解该名词进行提问，或者对该名词进行几句话的解释，然而，这就会造成受访者的压力，而且会影响其作答。如果要降低受访者这方面的作答压力，问卷可以设计成具有结构性质，比如对问题先进行分类，对每个问题都进行五等分评量，那么，即使要花费时间看问题，也不用花费时间看选项。然而，这个做法就会有随便回答、懒得回答、惯性作答以及作答催眠（越到后面，越不能集中精神）4 种降低信度的情况。

此外，访谈问题的回答方式尽量多样化。结构化问题能够保证各种超越个人、时间及研究者等来源的资料之间的可比性，因此在回答变量问题时特别有用，变量问题研究的是事情之间的差异。相反，非结构的方法允许研究者集中研究特定的现象，这些现象不同于其他现象，因此分别需要有针对性的方法。他们放弃普遍性与可比性，而追求内在效度与情境的理解，所以在揭示导致特定结果的过程方面非常有用。

5.6 如何有效回收问卷?

时间、地点、人力、礼品、经费和外部因素都是回收问卷要考虑的。但是回收问卷还更需要考虑方便性。能不能有效回收取决于问卷从受访者到研究者所需要的步骤和时间。例如：①邮寄问卷调查的回收，

包括放入回邮信封、粘好封口、询问信箱、找到信箱、投递，至少 6 个步骤。②电子问卷调查，则需要记得这回事、找到网址、(经作答后）点击完成，至少 3 个步骤。③面对面访谈，无论是约定时间回收，还是访谈中填写，都需要记得回收或访谈时间、找到研究者，至少两个步骤。④电话调查则不需要这些步骤。有趣的是，需要回收的步骤越多，相对地，允许问卷内的题目就越多。邮寄问卷的缺点，就是母体名册上没有正确的现居地址，而且回收的步骤对受访者而言也太麻烦。然而，相对于不能问太多问题的电话访谈、以及无法确定收信人的电子问卷（而且母体数据不提供电子邮件地址）而言，问卷调查适合采用邮寄问卷。而访谈调查，考虑到人数较少，采用面对面访谈。

5.7　如何追踪未回复的抽样样本?

理想的情况下，不需要追踪问卷，因为即使追踪到问卷，至少还需要进行一次假设检验：假设最近一次回收的问卷答案和前一次没有差别。在随机抽样的研究过程中，即使不需要追踪问卷，也要对“失踪问卷”和“无效问卷”进行分析，讨论是否具有影响抽样的失踪样本特征，亦即：是否有系统性误差的问题。

追踪问卷的方式和当初发放回收问卷的方式有关。不过，大体上不外乎通过电话、追踪邮递问卷、电子邮件、再次拜访访谈地点以及通过第一第二联络人间接转达这几种方式。这些无一例外涉及到母体名册的准确性。理论上，大家都相信通过多样而且不同的联系方式，总会“追踪”或者“找到”该位被抽样或者选中的未作答受访者。但是实际上，由于母体名册的不准确性，反而使得最后研究数据具有高度的“危险性”。

在《型人》研究之前，充分考虑了以下问题：①轻微错误的可能性：母体名册上充满了“虚假”信息，当时就有很多人不提供真实数据，所以抽样调查进行到一半，发现抽样无效，需要临时修改计划。②中度错误的可能性：后来通过上述许多方式，辗转联络到受访者，受访者知道研究者很“较真”所以给予回复，但是里面充满了涂涂改改的“认真”作答，事实上对研究者来说，增加了无效问卷率，甚至回答问卷的根本不是其本人。③重度错误的可能性：研究团队成员不考虑上述可能发生的统计错误，而把数据同其他数据一起编入数据库中，等到描述统计甚至推论统计都结束之后，由于理论模型的不稳定，导致重新拟合数据，结果发现数据库中出现大量的可能性错误数值。此时，需要重新进行数据清理，从头分析一遍，还要分析出错误率（对比两次分析的结果）。而最糟糕的是，如果先前没有对初始数据存档，那么连重做的机会都没有，研究者无法进行“侦错”与“除错”的工作，导致整个研究功败垂成。

5.8　如何有效降低抽样误差?

在《型人》研究中，采取了以下两种方式：①使各层的抽样比例，与母体中国科学院研究生院北京地区的系所结构比例维持均等（等比例)，使得样本问卷调查对象的构成可使母体构成做等比例抽选（比例抽样法）。因此，误差会比单纯的随机抽样少。②弹性地改变各层的抽样比例（不等比例)，因为中国科学院研究生院北京地区的各个系所规模大小不一，按照学生人数规模大小来改变抽样比例。规模大者，抽样比例就较高。这样也达到减少抽样误差的目的。

5.9　采用何种问卷发放与回收方式?

在问卷题项测量方面，《型人》以分类（category）尺度法来测定调查对象的态度、意见、评价、印象等心理层面的反应。 采用邮寄问卷法是基于在中国科学院研究生院北京地区进行调查的特殊情况：①因为调查对象多半忙于工作，不便或较多不易接受面谈法问卷调查，因此邮寄问卷可以节省他们的时间，不需要他们额外拿出时间与研究者联系和等待。②因为中国科学院研究生院北京地区各个系所的分布并不集中，这使得样本分布过度分散，唯有邮寄问卷法相对较为适合在这样的分布式学术机构中进行调查。③中国科学院的研究生是在各自相对独立的研究所中学习、研究或工作，因为中国科学院本身汇集了多所国家级重点实验室，出入门禁、安全检查标准、信息保密规定等，相当仔细与严格，因此如果要访问某一研究所时，程序相当繁琐，对调查多所研究所的本调查而言，用邮寄法来进行调查较为适当。④在缺乏经验丰富的调查员的情况下，邮寄法是减少人力又能获得可靠信息的有效方式之一。

5.10 如何避免无反应偏差？

邮寄法的回收率低，容易出现无反应偏差（non-response bias），换句话说，邮寄法的受访者对调查主题的兴趣会较容易左右整个调查结果。为了防止这种情况出现，当初研究采取了几个有效方法：①避免使用主题模糊不清的问卷，将主题定义为与研究生相关的题目“中国科学院研究生使用网络服务调查问卷”增加受访者参与感。②从“未回卷”的样本中随机抽样，进行当面访谈。③使用高回收率的固定样本（monitor sample）弥补实查阶段的误差。由于本次调查之前，研究者参与或者进行过 3 次不同时间与主题的问卷调查，因此有建立信任关系的研究生作为朋友，他们弥补了部分无反应偏差所可能造成的调查误差后果。

5.11 采用何种深度访谈？

当初因为需要在问卷调查的基础上，分析出不同行为的用户，试图从这些用户的访谈中，获得其行为的原因解释，所以采用深入面谈法。因为一般人无法接受长时间的面谈，但深度面谈法需要一段较长的时间，所以一方面需要从目的抽样中选取较大样本，另一方面也需要从这些样本的访谈意愿回馈中，识别适合进行访谈的对象，采取配额法来进行目的抽样。

5.12 如何确认问卷调查结果与访谈计划的紧密结合？

定性研究可以采用固定问题组访谈（nonscheduled-structured interview）。在《型人》中，调查问题是根据问卷调查结果而确定的。所以，访谈对象的选取是根据问卷调查分析结果所归类的 4 类用户群来确定的。不过，访谈问题的表述形式、顺序或回答方式等都没有严格规定，因此，允许访谈者用自己的语言和根据访谈具体环境来表述问题，允许访谈者根据访谈具体环境来决定问题顺序，允许被调查者用自己的语言和方式来回答问题，允许访谈者对不清楚的答案或可能有关联的其他问题作进一步的询问。但整个访谈要紧密围绕固定内容，所询问的各个问题不作实质变动。

在汇总资料方面，以表格化的方式整理。除汇总表之外，研究者将每一个受访者的回答反复精确，待对其生活及行为形成整体印象之后，再结合数据表格及事前假设等来进行分析。事实上，上述错误可以出现在任何一个研究步骤中，唯一能做的就是在研究过程中，详细地记录工作日志以及妥善地保管好每一次的数据文档，然后在研究结束后，仔细思考研究过程的细节。如果不遵守研究伦理，如果研究态度很随意，那么不可能确实完成整个研究计划。

6 合理推论原则

采用合理的竞替假设是探索性研究的主要方式之一。当研究者要开发新的研究领域时，常常借由探索性研究产生新的观点，然而它的缺点在于很少能圆满地回答问题。克服方法就是融会“假设验证”和“合理竞替”两套完全不同的思路，在《型人》中，因为考虑到探索性研究可以提供的是对于研究问题的提示以及研究方法的观察，以确切地回答研究问题，而不能提供绝对性的答案；所以，在理论建构的过程（目前还是假说阶段）里，尝试在后半段考虑方法、程序和人群的影响要素，判断调查结果与之前假设检验的对应关系。

在《型人》后半段中，采用的基本策略是“评估合理的竞替假设”，即检验和评估某特定现象是否存在着其他的替代性解释。此策略的适用范围不分其资料性质为量化或质化；不分研究方法的类型是实验法、横断法设计、纵贯法设计还是个案研究法；也不分资料收集方法是观察法、问卷法还是其他。因此在心态上，预期有其他诠释发现的方法，且视任何诠释在经过进一步检验后都是可修改的。不过，合理的竞替假设仍然有一定的推理限制与条件要求。

这种推理与它的“分析单位”息息相关，并且应该避免与分析单位有关的“区位谬误”和“化约主义”。分析单位可以划分为：个人、团体、组织以及人为事实，所以当《型人》分析单位是个人，即每一个用户，

而不是中国科学院或者工程热物理研究所等组织时，则没有因为根据某个研究所或团体的观察，来对个人作出结论的“区位谬误”的错误。同理反论，《型人》的研究出于邮寄问卷调查，回函也仅仅代表一部分人的情况，从这部分人当中选择受访者进行访谈所得出的结论，不适合推论于所有中国科学院研究生或者全中国人；所以，操作上又将它对应到《型人》前半段的假设检验中，通过反例举证原则，把网络用户行为的“常识”打破，得出可能具有差异化用户群体的情况，继而为后半段创建型人提供理论发挥的空间。在《型人》后半段中，先考虑到“化约主义”的影响（倾向于认定某一个分析单位或变项比其他分析单位或变项重要或相关，所以如果使用了不当的分析单位会造成谬误结论），所以采用合理的竞替假设，从不同的领域寻找不同的答案，在众多答案中，选择对研究结果（证据）最有可能的（likeliest）说明。因此，合理的竞替假设不仅仅包括了来自图书情报学的假设（如信息搜寻行为），也来自于其他领域（如心理学、社会学、经济学、人类学、教育学等）的假设。

换句话说，“合理竞替”不是按照自己的意思随意解释，因为“合理性”不是来自研究者怎么方便就怎么想，怎么想就怎么写，怎么写就怎么发表，而是“竞争”和“替代”的结果。学术理论并非真理，只是相对合理性，而这种合理性来自辩证和竞替。如此一来，反复苛求的结果就是《型人》研究价值与贡献的最后呈现，即完成一套能够验证和简单操作的假说。

7 工具：关于 SPSS v12 的操作

研究没有捷径，唯有多思考、多尝试、多实战。不过，按图索骥可以节省部分摸索时间，以下是整理自相关教科书，并且在攻读博士期间（2006-2009 年）实际操作过 SPSS v12 的笔记。如前所述，工具操作一点也不难，按部就班，即可实现数据和图表产出。关键是研究伦理与态度：如果立意良善，即便是简单描述统计，也能产出令人尊敬的研究；同理，如果一味强调应用了绚丽的数学方程式而忽略了思考朴质的研究伦理，则漂亮图表下的解释内容必定破洞百出。

- 简单描述统计

  ①打开文件

  ②Statistics→ Summarize → Descriptives

  ③选择变项，按“←”Variable(s)

  ④Options → 选 Statistics 的所有选项 → Continue

  ⑤OK

  Mean 算术平均值

  Variance 方差

  S.E.mean 平均数的标准误差

  Skewness 偏度

  Range 最大至于最小值的差

  Kurtosis 峰度

  Sum 总和

  Minimum 最小值

  Maximum 最大值

  Std.deviation 标准差

- b. 如何知道定类变项和定类变项有无关系？——卡方分析(不用 SPSS)

  ①与假设相反的零假设

  ②自由度 df=(r -1)(c-1)

r 行的栏目数；c 列的栏目数
③根据 df 和显著水平（0.01 或 0.05 或 0.1）查“卡方分布表”
④得到“检验参数“
⑤利用公式
⑥得到“卡方值”
⑦比较“检验参数”和“卡方值”
差距过大，否认零假设，从而假设成立
无显著差异，支持零假设，从而假设不成立

- 如何知道定类变项和定距变项有无关系？——单因素方差分析

①打开文件
②Statistics→ Compare Means → One-Way ANOVA
③选择因变项，按“→”Dependent List
④选择自变量，按“→”Factor
⑤OK
差距过大，否认零假设，从而假设成立
无显著差异，支持零假设，从而假设不成立

- 如何知道定距变项和定距变项有无关系？——利用皮尔逊相关系数——R 系数

①打开文件
②Statistics→ Correlate → Bivariate
③选择要分析的变项，按“→”Variables
④Correlation Coefficients → Pearson
⑤Test of Significance → Two-tailed
⑥Options → 选 Statistics 的所有选项 → Continue
⑦OK
0<[R]<0.3　无相关
0.3<[R]<0.5 低度相关
0.5<[R]<0.7 中度相关
0.7<[R]<1　高度相关（可以进行一元回归）

- 如何测量两个变项的变化趋势？——一元回归

①打开文件
②Statistics→ Regression → Liner
③选择因变项，按“→”Dependent
④选择自变项，按“→”Independent(s)
⑤Method → Stepwise
⑥Standardized Residual Plots →Histogram
⑦OK

- 如何描述对一个个体有关的多个变项（变项之间不相关）——因子分析、聚类分析

①打开文件
②Statistics→ Data Reduction → Factor
③Descriptives→ 选 Univariate descriptives、Initial solution、Coefficients、Significance
④Extraction → Principal components
⑤Extract → Eigenvalues over →“>1”

⑥Display → 选 Unrotated factor solution、Scree plot
⑦Factor Scores → Save as variables
⑧Method → Regression、Display factor score coefficient matrix
⑨OK

- 如何描述对一个个体有关的多个变项（变项之间不相关）——聚类分析

② 在因子分析后，获得“fac1-1”和“fac2-1”
②Statistics→ Clussify→ Hierarchical Cluster → 选“fac1-1”和“fac2-1”
③Plot→ Dengrogram、All clusters、Vertical→ Continue
④Method→ Between-groups linkage
⑤Save New Variables→ Single solution 3 clusters
⑥OK

用户行为研究的经验，来自博士研究期间的导师、老师们和朋友们，他们对研究者产生许多直接和间接的影响。相关指导性图书，可以参考张晓林《信息管理学研究方法》、Earl Babbie《社会科学研究方法》等。在 3 年中，进行 6 次大规模社会科学调查研究的经验，虽然只限于位于北京市的中国科学院，但是对笔者日后在香港城市大学支援经济金融系的工作、在“国立”台湾大学的机构排名与科学家行为研究以及为本土公司和跨国企业开发市场客户群等工作都有帮助。

虽然终端用户在各地区各个机构的属性不同、特征不同、行为不同，而且服务于不同的工作岗位，他们的立场各有不同，但是因为对用户行为的敏感度高，便能够快速进入研究前沿或者突破市场障碍，在各种职场上贡献自己的能力，为图书情报事业尽一份绵薄之力。抛砖引玉，是以为文。

[作者简介] 顾立平，男，1978 年生，博士，发表论文 33 篇。